

ARBITRATION SYSTEM AND METHOD FOR MEMORY RESPONSES IN A
HUB-BASED MEMORY SYSTEM

TECHNICAL FIELD

This invention relates to computer systems, and, more particularly, to a
5 computer system including a system memory having a memory hub architecture.

BACKGROUND OF THE INVENTION

Computer systems use memory devices, such as dynamic random access memory (“DRAM”) devices, to store data that are accessed by a processor. These memory devices are normally used as system memory in a computer system. In a
10 typical computer system, the processor communicates with the system memory through a processor bus and a memory controller. The processor issues a memory request, which includes a memory command, such as a read command, and an address designating the location from which data or instructions are to be read. The memory controller uses the command and address to generate appropriate command signals as
15 well as row and column addresses, which are applied to the system memory. In response to the commands and addresses, data are transferred between the system memory and the processor. The memory controller is often part of a system controller, which also includes bus bridge circuitry for coupling the processor bus to an expansion bus, such as a PCI bus.

20 Although the operating speed of memory devices has continuously increased, this increase in operating speed has not kept pace with increases in the operating speed of processors. Even slower has been the increase in operating speed of memory controllers coupling processors to memory devices. The relatively slow speed of memory controllers and memory devices limits the data bandwidth between the
25 processor and the memory devices.

In addition to the limited bandwidth between processors and memory devices, the performance of computer systems is also limited by latency problems that increase the time required to read data from system memory devices. More specifically,

when a memory device read command is coupled to a system memory device, such as a synchronous DRAM (“SDRAM”) device, the read data are output from the SDRAM device only after a delay of several clock periods. Therefore, although SDRAM devices can synchronously output burst data at a high data rate, the delay in initially providing
5 the data can significantly slow the operating speed of a computer system using such SDRAM devices.

One approach to alleviating the memory latency problem is to use multiple memory devices coupled to the processor through a memory hub. In a memory hub architecture, a memory hub controller is coupled over a high speed data link to
10 several memory modules. Typically, the memory modules are coupled in a point-to-point or daisy chain architecture such that the memory modules are connected one to another in series. Thus, the memory hub controller is coupled to a first memory module over a first high speed data link, with the first memory module connected to a second memory module through a second high speed data link, and the second memory module
15 coupled to a third memory module through a third high speed data link, and so on in a daisy chain fashion.

Each memory module includes a memory hub that is coupled to the corresponding high speed data links and a number of memory devices on the module, with the memory hubs efficiently routing memory requests and memory responses
20 between the controller and the memory devices over the high speed data links. Each memory request typically includes a memory command specifying the type of memory access (*e.g.*, a read or a write) called for by the request, a memory address specifying a memory location that is to be accessed, and, in the case of a write memory request, write data. The memory request also normally includes information identifying the memory
25 module that is being accessed, but this can be accomplished by mapping different addresses to different memory modules. A memory response is typically provided only for a read memory request, and typically includes read data as well as an identifying header that allows the memory hub controller to identify the memory request corresponding to the memory response. However, it should be understood that memory
30 requests and memory responses having other characteristics may be used. In any case,

in the following description, memory requests issued by the memory hub controller propagate downstream from one memory hub to another, while memory responses propagate upstream from one memory hub to another until reaching the memory hub controller. Computer systems employing this architecture can have a higher bandwidth
5 because a processor can access one memory device while another memory device is responding to a prior memory access. For example, the processor can output write data to one of the memory devices in the system while another memory device in the system is preparing to provide read data to the processor. Moreover, this architecture also provides for easy expansion of the system memory without concern for degradation in
10 signal quality as more memory modules are added, such as occurs in conventional multi drop bus architectures.

Although computer systems using memory hubs may provide superior performance, they nevertheless may often fail to operate at optimum speeds for a variety of reasons. For example, even though memory hubs can provide computer systems with
15 a greater memory bandwidth, they still suffer from latency problems of the type described above. More specifically, although the processor may communicate with one memory device while another memory device is preparing to transfer data, it is sometimes necessary to receive data from one memory device before the data from another memory device can be used. In the event data must be received from one
20 memory device before data received from another memory device can be used, the latency problem continues to slow the operating speed of such computer systems.

Another factor that can reduce the speed of memory transfers in a memory hub system is the transferring of read data upstream (*i.e.*, back to the memory hub controller) over the high-speed links from one hub to another. Each hub must
25 determine whether to send local responses first or to forward responses from downstream memory hubs first, and the way in which this is done affects the actual latency of a specific response, and more so, the overall latency of the system memory. This determination may be referred to as arbitration, with each hub arbitrating between local requests and upstream data transfers.

There is a need for a system and method for arbitrating data transfers in a system memory having a memory hub architecture to lower the latency of the system memory.

SUMMARY OF THE INVENTION

5 According to one aspect of the present invention, a memory hub includes a local queue that receives and stores local memory responses. A bypass path receives downstream memory responses and passes the downstream memory responses while a buffered queue is coupled to the bypass path and stores downstream memory responses. A multiplexer is coupled to the local queue, the bypass path, and the buffered queue,
10 and outputs one of the responses responsive to a control signal. Arbitration control logic is coupled to the multiplexer and develops the control signal to control the source of the responses output by the multiplexer.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is a block diagram of a computer system including a system
15 memory having a high bandwidth memory hub architecture according to one example of the present invention.

Figure 2 is a functional block diagram illustrating an arbitration control component contained in each of the memory hubs of Figure 1 according to one example of the present invention.

20 Figure 3 is a functional flow diagram illustrating the flow of upstream memory responses in a process executed by the arbitration control component of Figure 2 where downstream responses are give priority over local responses according to one embodiment of the present invention.

Figure 4 is a functional flow diagram illustrating the flow of upstream
25 memory responses in a process executed by the arbitration control component of Figure 2 to provide equal bandwidth for local and downstream memory responses.

DETAILED DESCRIPTION OF THE INVENTION

A computer system 100 according to one example of the present invention is shown in Figure 1. The computer system 100 includes a system memory 102 having a memory hub architecture including a plurality of memory modules 130, 5 each memory module including a corresponding memory hub 140. Each of the memory hubs 140 arbitrates between memory responses from the memory module 130 on which the hub is contained and memory responses from downstream memory modules, and in this way the memory hubs effectively control the latency of respective memory modules in the system memory by controlling how quickly responses are returned to a system 10 controller 110, as will be described in more detail below. In the following description, certain details are set forth to provide a sufficient understanding of the present invention. One skilled in the art will understand, however, that the invention may be practiced without these particular details. In other instances, well-known circuits, 15 control signals, timing protocols, and/or software operations have not been shown in detail or omitted entirely in order to avoid unnecessarily obscuring the present invention.

The computer system 100 includes a processor 104 for performing various computing functions, such as executing specific software to perform specific calculations or tasks. The processor 104 is typically a central processing unit (“CPU”) 20 having a processor bus 106 that normally includes an address bus, a control bus, and a data bus. The processor bus 106 is typically coupled to cache memory 108, which, as previously mentioned, is usually static random access memory (“SRAM”). Finally, the processor bus 106 is coupled to the system controller 110, which is also sometimes referred to as a “North Bridge” or “memory controller.”

25 The system controller 110 serves as a communications path to the processor 104 for the memory modules 130 and for a variety of other components. More specifically, the system controller 110 includes a graphics port that is typically coupled to a graphics controller 112, which is, in turn, coupled to a video terminal 114. The system controller 110 is also coupled to one or more input devices 118, such as a 30 keyboard or a mouse, to allow an operator to interface with the computer system 100.

Typically, the computer system 100 also includes one or more output devices 120, such as a printer, coupled to the processor 104 through the system controller 110. One or more data storage devices 124 are also typically coupled to the processor 104 through the system controller 110 to allow the processor 104 to store data or retrieve data from 5 internal or external storage media (not shown). Examples of typical storage devices 124 include hard and floppy disks, tape cassettes, and compact disk read-only memories (CD-ROMs).

The system controller 110 also includes a memory hub controller (“MHC”) 132 that is coupled to the system memory 102 including the memory modules 10 130a,b...n, and operates to apply commands to control and access data in the memory modules. The memory modules 130 are coupled in a point-to-point or daisy chain architecture through respective high speed links 134 coupled between the modules and the memory hub controller 132. The high-speed links 134 may be optical, RF, or electrical communications paths, or may be some other suitable type of communications 15 paths, as will be appreciated by those skilled in the art. In the event the high-speed links 134 are implemented as optical communications paths, each optical communication path may be in the form of one or more optical fibers, for example. In such a system, the memory hub controller 132 and the memory modules 130 will each include an optical input/output port or separate input and output ports coupled to the corresponding 20 optical communications paths. Although the memory modules 130 are shown coupled to the memory hub controller 132 in a daisy architecture, other topologies that may be used, such as a ring topology, will be apparent to those skilled in the art.

Each of the memory modules 130 includes the memory hub 140 for communicating over the corresponding high-speed links 134 and for controlling access 25 to six memory devices 148, which are synchronous dynamic random access memory (“SDRAM”) devices in the example of Figure 1. The memory hubs 140 each include input and output ports that are coupled to the corresponding high-speed links 134, with the nature and number of ports depending on the characteristics of the high-speed links. A fewer or greater number of memory devices 148 may be used, and memory devices 30 other than SDRAM devices may also be used. The memory hub 140 is coupled to each

of the system memory devices 148 through a bus system 150, which normally includes a control bus, an address bus, and a data bus.

As previously mentioned, each of the memory hubs 140 executes an arbitration process that controls the way in which memory responses associated with the 5 memory module 130 containing that hub and memory responses from downstream memory modules are returned to the memory hub controller 132. In the following description, upstream memory responses associated with the particular memory hub 140 and the corresponding memory module 130 will be referred to as "local" upstream memory responses or simply "local responses," while upstream memory responses from 10 downstream memory modules will be referred to as downstream memory responses or simply "downstream responses." In operation, each memory hub 140 executes a desired arbitration process to control the way in which local and downstream responses are returned to the memory hub controller 132. For example, each hub 140 may give priority to downstream responses and thereby forward such downstream responses 15 upstream prior to local responses that need to be sent upstream. Conversely, each memory hub 140 may give priority to local responses and thereby forward such local responses upstream prior to downstream responses that need to be sent upstream. Examples of arbitration processes that may be executed by the memory hubs 140 will be described in more detail below.

20 Each memory hub 140 may execute a different arbitration process or all the hubs may execute the same process, with this determination depending on the desired characteristics of the system memory 102. It should be noted that the arbitration process executed by each memory hub 140 is only applied when a conflict exists between local and downstream memory responses. Thus, each memory hub 140 need 25 only execute the corresponding arbitration process when both local and downstream memory responses need to be returned upstream.

Figure 2 is a functional block diagram illustrating an arbitration control component 200 contained in the memory hubs 140 of Figure 1 according to one embodiment of the present invention. The arbitration control component 200 includes 30 two queues for storing associated memory responses. A local queue 202 receives and

stores local memory responses LMR from the memory devices 140 on the associated memory module 130. A buffered queue 206 receives and stores downstream memory responses which cannot be immediately forwarded upstream through a bypass path 204. A multiplexer 208 selects responses from one of the queues 202, 206 or the bypass path 5 204 under control of arbitration control logic 210 and supplies the memory responses in the selected queue upstream over the corresponding high-speed link 134. The arbitration control logic 210 is coupled to the queues 202, 206 through a control/status bus 136, which allows the logic 210 to monitor the contents of each of the queues 202, 206, and utilizes this information in controlling the multiplexer 208 to thereby control 10 the overall arbitration process executed by the memory hub 140. The control/status bus 136 also allows “handshaking” signals to be coupled from the queues 202, 206 to the arbitration logic 210 to coordinate the transfer of control signals from the arbitration logic 210 to the queues 202, 206.

The specific operation of the arbitration control logic 210 in controlling 15 the multiplexer 208 to provide responses from one of the queues 202, 206 or the bypass path 204 depends on the particular arbitration process being executed by the control logic. Several example arbitration processes that may be executed by the control logic 210 will now be described in more detail with reference to Figures 3 and 4. Figure 3 is a functional flow diagram illustrating the flow of upstream memory responses in a 20 process executed by the arbitration control component 200 of Figure 2 where downstream responses are given priority over local responses according to one embodiment of the present invention. In the example of Figure 3, the memory hub controller 132 applies a memory request to each of the memory modules 130a, 130b, and 130c. Each of the memory modules 130a-c provides a corresponding upstream 25 response in response to the applied request, with the responses for the modules 130a, 130b, and 130c being designated A1, B1, and C1, respectively. The responses B1 and C1 are assumed to arrive at the local queue 202 and bypass path 204 in the hub 140 of the module 130b at approximately the same time. In this embodiment, the arbitration control logic 210 gives priority to downstream responses, and as a result the hub 140 in

module 130b forwards upstream the downstream responses C1 first and thereafter forwards upstream the local response B1 as shown in Figure 3.

If the response C1 arrives in the bypass path 204 in the hub 140 of the module 130a at approximately the same time as the local response A1 arrives in the local queue 202, the arbitration control logic 210 forwards upstream the downstream response C1 prior to the local response A1. Moreover, if the response B1 arrives in the bypass path 204 in the hub 140 of module 130a at approximately the same time as the downstream response C1, then arbitration control logic 210 forwards upstream the downstream response C1 followed by response B1 followed by local response A1, as shown in Figure 3. The system controller 110 thus receives the responses C1, B1, and A1 in that order.

Because the arbitration control logic 210 in each memory hub 140 may execute an independent arbitration process, the arbitration control logic in the memory hub of the module 130a could give priority to local responses over downstream responses. In this situation, if the responses C1 and B1 arrive at the bypass path 204 in the hub 140 of the module 130a at approximately the same time as the local response A1 arrives in the local queue 202, the arbitration control logic 210 forwards upstream the local response A1 prior to the downstream responses C1 and B1. The memory hub controller 132 thus receives the responses A1, C1 and B1 in that order, as shown in parentheses in Figure 3. Thus, by assigning different arbitration processes to different memory hubs 140 the latency of the corresponding memory modules 130 may be controlled. For example, in the first example of Figure 3 where priority is given to downstream responses, the latency of the module 130a is higher than in the second example where in module 130a priority is given to local responses. In the second example, the memory hub controller 132 could utilize the module 130a to store frequently accessed data so that the system controller can more quickly access this data. Note that in the second example the responses C1, B1 would first be transferred to the buffered queue 206 since they could not be forwarded upstream immediately, and after response A1 is forwarded the responses C1, B1 would be forwarded from the buffered queue.

Figure 4 is a functional flow diagram illustrating the flow of upstream memory responses in a process executed by the arbitration control component 200 of Figure 2 to alternate between a predetermined number of responses from local and downstream memory . In the example of Figure 4, the memory hub controller 132 applies two memory requests to each of the memory modules 130a, 130b, and 130c, with the requests applied to module 130a being designated A1, A2, requests applied to module 130b being designated B1, B2, and requests to module 130c being designated C1, C2. The responses C1 and C2 are assumed to arrive at the bypass path 204 in the hub 140 of the module 130b at approximately the same time as the local responses B1,
5 B2 arrive at the local queue 202. The responses C1, C2 are transferred to the buffered queue 206 since they cannot be forwarded upstream immediately. The arbitration control logic 210 thereafter alternately forwards responses from the local queue 202 and the buffered queue 206. In the example of Figure 4, the local response B1 from the local queue 202 is forwarded first, followed by the downstream response C1 from the buffered queue 206, then the local response B2 and finally the downstream response C2.
10
15

Now in the module 130a, the responses B1, C1, B2, C2 are assumed to arrive at the bypass path 204 in the hub 140 at approximately the same time as the local responses A1, A2 arrive at the local queue 202. The responses B1, C1, B2, C2 are transferred to the buffered queue 206 since they cannot be forwarded upstream immediately. The arbitration control logic 210 thereafter operates in the same way to alternately forward responses from the local queue 202 and the buffered queue 206. The local response A1 from the local queue 202 is forwarded first, followed by the downstream response B1 from the buffered queue 206, then the local response A2 followed by downstream response C1. At this point, the local queue 202 is empty while
20
25 the buffered queue 206 still contains the responses B2, C2. No conflict between local and downstream responses exists, and the arbitration control logic 200 accordingly forwards upstream the remaining responses B2, C2 to empty the buffered queue 206.

In the arbitration process illustrated by Figure 4, the arbitration control logic 210 forwarded a predetermined number of either local or downstream responses prior to forwarding the other type of response. For example, in the process just
30

described the arbitration control logic 210 forwards one local response and then one downstream response. Alternatively, the arbitration control logic 210 could forward two local responses followed by two downstream responses, or three local responses followed by three downstream responses, and so on. Furthermore, the arbitration
5 control logic 210 could forward N local responses followed by M downstream responses, where N and M may be selected to give either local or downstream responses priority.

In another embodiment, the arbitration control logic 210 of Figure 2 executes an oldest first algorithm in arbitrating between local and downstream memory
10 responses. In this embodiment, each memory response includes a response identifier portion and a data payload portion. The response identifier portion identifies a particular memory response and enables the arbitration control logic 210 to determine the age of a particular memory response. The data payload portion includes data being forwarded upstream to the memory hub controller 132, such as read data. In operation,
15 the arbitration control logic 210 monitors the response identifier portions of the memory responses stored in the local queue 202 and the buffered queue 206 and selects the oldest response contained in either of these queues as the next response to be forwarded upstream. Thus, independent of queue 202, 206 in which a memory response is stored, the arbitration control logic 210 forwards the oldest responses first.

20 In determining the oldest response, the arbitration control logic 210 utilizes the response identifier portion of the memory response and a time stamp assigned to the memory request corresponding to the response. More specifically, the memory hub controller 132 generates a memory request identifier for each memory request. As the memory request passes through each memory hub 140, the arbitration
25 control logic 210 of each hub assigns a time stamp to each request, with the time stamp indicating when the request passed through the memory hub 140. Thus, each hub 140 essentially creates a table of request identifiers and associated time stamps. Thus, the control logic 210 in each hub 140 stores a table of a unique memory request identifier and a corresponding time stamp for each memory request passing through the hub.

In each memory response, the response identifier portion corresponds to the memory request identifier, and thus the response for a given a request is identified by the same identifier. The arbitration control logic 210 thus identifies each memory response stored in the local queue 202 and buffered queue 206 by the corresponding 5 response identifier portion. The control logic 210 then compares the response identifier portion of each response in the queues 202, 206 to the table of request identifiers, and identifies the time stamp of the response identifier as the time stamp associated with the corresponding request identifier in the table. The control logic 210 does this for each 10 response, and then forwards upstream the oldest response as indicated by the corresponding time stamp. The arbitration control logic 210 repeats this process to determine the next oldest response and then forwards that response upstream, and so on.

In the preceding description, certain details were set forth to provide a sufficient understanding of the present invention. One skilled in the art will appreciate, however, that the invention may be practiced without these particular details. 15 Furthermore, one skilled in the art will appreciate that the example embodiments described above do not limit the scope of the present invention, and will also understand that various equivalent embodiments or combinations of the disclosed example embodiments are within the scope of the present invention. Illustrative examples set forth above are intended only to further illustrate certain details of the various 20 embodiments, and should not be interpreted as limiting the scope of the present invention. Also, in the description above the operation of well known components has not been shown or described in detail to avoid unnecessarily obscuring the present invention. Finally, the invention is to be limited only by the appended claims, and is not limited to the described examples or embodiments of the invention.